

# The acoustic features of vowel-like *grunt* calls in chacma baboons (*Papio cyncephalus ursinus*): Implications for production processes and functions

Michael J. Owren

Department of Psychology, Reed College, 3203 SE Woodstock Boulevard, Portland, Oregon 97202

Robert M. Seyfarth and Dorothy L. Cheney

Departments of Psychology and Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 9 July 1996; accepted for publication 26 December 1996)

The acoustic features of 216 baboon *grunts* were investigated through analysis of field-recorded calls produced by identified females in known contexts. Analyses addressed two distinct questions: whether the acoustic features of these tonal sounds could be characterized using a source-filter approach and whether the acoustic features of *grunts* varied by individual caller and social context. Converging evidence indicated that *grunts* were produced through a combination of periodic laryngeal vibration and a stable vocal tract filter. Their acoustic properties closely resembled those of prototypical human vowel sounds. In general, variation in the acoustic features of the *grunts* was more strongly related to caller identity than to the social contexts of calling. However, two acoustic parameters, second formant frequency and overall spectral tilt, did vary consistently depending on whether the caller was interacting with an infant or participating in a group move. Nonetheless, in accordance with the general view that identity cueing is a compelling function in animal communication, it can be concluded that much of the observed variability in *grunt* acoustics is likely to be related to this aspect of signaling. Further, cues related to vocal tract filtering appear particularly likely to play an important role in identifying individual calling animals. © 1997 Acoustical Society of America. [S0001-4966(97)01605-6]

PACS numbers: 43.80.Ka, 43.70.Aj [FD]

## INTRODUCTION

Studies of nonhuman primates (hereafter *primates*) have demonstrated many essential anatomical similarities in the sound-production mechanisms of these animals and humans (e.g., Geist, 1933; Negus, 1949; Hill and Booth, 1957; Kelemen, 1969; Hilloowala, 1975; Schon Ybarra, 1995). Because many primate vocalizations, like human speech sounds, are produced using vocal-fold vibration and one or more contiguous supralaryngeal vocal tract chambers (see Fig. 1 for schematic representations of the human and baboon vocal tract, respectively), researchers have long noted that these sounds are necessarily formed through the interaction of source energy input with resonances (*formants*) due to vocal tract cavities (e.g., Kelemen, 1938, 1949; Negus, 1949; Zhinkin, 1963). These concepts, well-developed in speech applications (e.g., Chiba and Kajiyama, 1941; Stevens and House, 1955; Fant, 1960), have significantly influenced both analysis and synthesis of speech [see Rubin and Vatikiotis-Bateson (in press) for a recent review]. In contrast, with the notable exception of work by Lieberman and his colleagues (e.g., Lieberman, 1968, 1969; Lieberman *et al.*, 1969, 1972) and Zhinkin's (1963) largely unacknowledged work, production-based perspectives have historically not influenced acoustic analysis of primate calls. More recently, however, both empirical and tutorial papers have appeared that argue for the application of source-filter theory and other speech-related, production-based perspectives in acoustic primatology (see reviews by Fitch and Hauser, 1995; Owren and Linker, 1995).

Briefly summarizing the available evidence concerning the contribution of vocal-fold-related source energy in primates, many species appear to employ both humanlike vocal-fold action and vibration modes that humans can probably not emulate. For instance, a number of species can produce highly tonal vocalizations involving stable fundamental frequencies and rich harmonic spectra [e.g., baboons (Andrew, 1976); Japanese macaques (Green, 1975); gelada monkeys (Richman, 1976); rhesus macaques (Hauser *et al.*, 1993)]. Fundamentals in these and more pure-tone vocalizations show both significant intra- and inter-species variation, for instance, being somewhat lower [e.g., baboons (Owren *et al.*, 1993)], moderately higher [e.g., rhesus and Japanese macaques (Owren *et al.*, 1992)], or many times higher [e.g., bushbabies and slow loris (Zimmermann, 1981); chimpanzees (Clark and Wrangham, 1993); and rhesus macaques (Hauser *et al.*, 1993)] than corresponding values in human phonation.

However, while laryngographic measurements by Brown and colleagues (personal communication) showed that normative variability in the periodicity of vocal-fold vibration were quite similar in a Sykes's monkey and a human subject, macaques and other species also routinely produce calls based on vibration modes that are quite unlike those underlying human speech sounds. Such vocalizations can include high-amplitude, noise-based sounds, virtually pure-tone sinusoids, frequency sweeps that can cover many octaves in a fraction of a second, and apparent simultaneous combinations of periodic and aperiodic vocal fold vibration

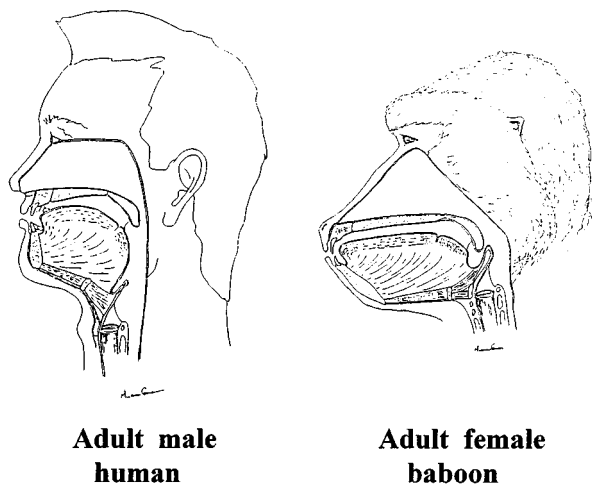


FIG. 1. Schematic, midsagittal views of a human and a baboon vocal tract. The relative headsize shown is approximately correct, although the drawings were not created to exact specifications. From Owren and Linker (1995), adapted with permission.

patterns [see, for example, spectrograms shown in Peters (1983) and Green (1975) for rhesus and Japanese macaques, respectively]. Anatomical studies by Schön Ybarra (1995) and others have indicated a probable basis for these capabilities in the occurrence of a relatively rigid “lip” on the medial extremity of each vocal fold in a number of primate species.

Sound spectrography has also revealed strong evidence of nonlaryngeal influences on vocalizations in many primates. Some, for instance, have been found to have extralaryngeal vocal sacs proposed to act as resonators in sounds specialized for long-distance transmission (see Gautier and Gautier, 1977). More commonly, however, spectrographic evidence shows apparent resonance effects that are analogous to those of the human supralaryngeal vocal tract, for instance in both tonal *coo* calls and other vocalizations in macaques (e.g., Grimm, 1967; Lieberman, 1968; Green, 1975; Peters, 1983; Hohmann and Herzog, 1985; Hohmann, 1989; Masataka and Thierry, 1993), in gelada monkey calls (Richman, 1976), in *double-grunts* produced by wild mountain gorillas (Seyfarth *et al.*, 1994), and in the baboon *grunts* (Andrew, 1976) that are the subject of this article.

Modification of vocal tract resonances has also been both proposed (e.g., Lieberman, 1975; Richman, 1976; Hauser and Fowler, 1991; Hauser, 1992) and indirectly demonstrated (e.g., Hauser *et al.*, 1993; Hauser and Schön Ybarra, 1994; Brown and Cannito, 1995). Such modification involves changes not only in laryngeal, lip, and mandible positioning, but may also include tongue movement. As illustrated in Fig. 1, primates differ from humans in having thinner tongues, larynges that are positioned higher in the neck, and a relative lack of soft tissues in the supralaryngeal cavities. Thus the former appear to have significantly less opportunity to alter the resonance properties of their vocal tract through flexible modification of articulator position than do the latter. While the functional significance of the spectral changes that can occur has not been tested, the perceptual salience of static spectral patterning has been demonstrated

both in vervets classifying synthetic versions of noisy species-typical alarm calls (Owren, 1990) and in baboons and Japanese macaques discriminating among human vowel sounds (Hienz and Brady, 1988; Sinnott, 1989).

Overall then, it is clear that the source-filter perspective can be usefully applied to a number of calls produced by primates, and that both vocal-fold action and extra-laryngeal resonance may play significant roles in shaping the functionally important acoustic characteristics of such calls. On the other hand, insufficient evidence is available to specify many details of either the source or filter components of production for any given call in any particular primate species, or concerning possible interaction effects between the two. Calls produced by baboons (*Papio cynocephalus*) may be of particular interest in this regard, due to the resemblance of these calls to human vowel sounds (Zhinkin, 1963; Andrew, 1976; Owren *et al.*, 1993, 1995). These similarities set the stage for detailed examination of grunts both for evidence of basic sound-producing processes and for the role that particular acoustic features, whether speechlike or not, play in the use of these sounds as species-typical social signals. In general, although a voluminous body of work exists on baboon social behavior, ecology, and reproduction, relatively little is known concerning either the acoustics or social functions of baboon calls. However, a brief description of the baboon vocal repertoire is provided by Bolwig (1959), while more detailed analyses of the acoustics and functions of grunts (Zhinkin, 1963; Andrew, 1976; Waser, 1982; Cheney *et al.*, 1995a), *contact barks* (Byrne, 1981; Cheney *et al.*, 1996), *copulatory* calls (Hamilton and Arrowood, 1978; O’Connell and Cowlshaw, 1994), and *loud calls* (Waser, 1982; Brown *et al.*, 1995) are also available.

The work reported here is an acoustic analysis of a sample of grunts recorded from a group of chacma baboons (*P. c. ursinus*) living under natural conditions in Botswana. There were two overriding goals for these analyses. First, we wanted to examine as well as possible the apparent contributions of laryngeally based sound generation and vocal tract filtering in the calls. To do so, we first characterized the apparent fundamental frequency of each vocalization through three different methods: (1) automated pitch extraction from the time waveform; (2) measuring the lowest visible harmonic in spectrographic representations; and (3) ascertaining the frequency intervals between partials that appeared above the apparent fundamental frequency. Then we examined the overall patterning of the frequency spectrum for evidence of formant-related spectral energy peaks and compared the relationship between derived values for fundamental frequencies and formant frequencies in each call. Finally, we used formant frequency measurement to estimate the baboon vocal tract length (based on an idealized, uniform-tube model) for comparison to values from humans.

The second overriding goal was to examine possible acoustic differences between grunts produced in varying behavioral circumstances, paying particular attention to the possibility that baboons might produce calls with varying spectral properties through apparent flexible modification of articulator positioning in the supralaryngeal vocal tract. As reported by Cheney *et al.* (1995a), baboons produce grunts

under a variety of social conditions, including situations in which these vocalizations appear to be directed toward particular receivers and act to facilitate further social interaction. As a first test of the relation between acoustic features and particular behavioral contexts, calls were scored according to the identity and relative social rank of the callers and apparent receivers (specific receivers were not always evident), whether an infant was involved, and whether the vocalization appeared as a component of initiating or participating in group movement. A variety of acoustic measures were made from each call and mean values for each were compared according to these context variables. Discriminant function analysis was then used to measure the classification accuracy achieved using each of these acoustic and social contexts.

## I. METHOD

### A. Study area and subjects

The study site was in the Okavango Delta, a vast seasonal swamp created by the floodplain of the Okavango River in northwestern Botswana. The habitat consisted of flat, open grasslands interspersed by slightly elevated wooded patches. Subjects were nine sexually mature, adult females that were part of a group of approximately 70 baboons. The group has been observed continuously since 1977 by Hamilton and colleagues (e.g., Bulger and Hamilton, 1988; Hamilton and Bulger, 1992) and ourselves (e.g., Cheney *et al.*, 1995a, b). All animals can be identified individually and maternal relatedness is known for natal individuals. Relative dominance ranks are also known for each animal. The group is fully habituated to the presence of human observers on foot.

### B. Observational protocol and behavioral-context definitions

Data collection occurred continuously between July 1992 and July 1993, beginning each day as the group moved from the sleeping site in the morning and continuing for a predetermined interval of up to two and one-half hours. During data collection, two to four observers spread themselves throughout the group and noted the identities of as many calling females as possible. For each calling event or bout, the apparent receiver and salient behaviors by any individual occurring before and after the calling event were also noted. Other potentially significant circumstances (e.g., whether another group was nearby or whether a predator had recently been seen) were also available in the call records, allowing us to restrict the sample unambiguously to the variables of greatest interest.

Three primary variables were examined in detail in testing for possible differentiation in grunt call acoustics. First, we checked for individual variation in the parameters measured in each grunt based on the identity of the calling animal (variable *Caller*). Second, the behavioral context (variable *Context*) associated with each grunt was noted according to whether the call was produced when the calling

individual (and often others as well) were moving from one area of their range to another (*Move*) or whether the calling individual was engaged in interactions with an infant or the infant's mother (*Infant*). The *Move* and *Infant* contexts represented over 80% of all circumstances in which grunts were recorded. Both prior research on vervet monkeys (Cheney and Seyfarth, 1982; Seyfarth and Cheney, 1984) and impressions from field observations led us to hypothesize that, if the acoustic features of grunts were differentially associated with various social contexts, these two particular variables were the most likely to reveal such relationships. Other analyses have examined grunts given when animals are reconciling with one another following an agonistic interaction (Cheney *et al.*, 1995; Cheney and Seyfarth, in press). These grunts are also relatively common, and appear to be acoustically indistinguishable from those classified as *Infant* grunts in the current analyses (Seyfarth and Cheney, 1997).

The identity of the receiver was also recorded whenever evident, and calls to females were scored as occurring either to subordinate or to dominant animals (variable *Receiver-Rank*). This variable was also examined in separate analyses, due to the evident importance of relative dominance rank in many nonhuman primate species, including baboons (e.g., Smuts *et al.*, 1987), and previous evidence from other species that calls to dominant animals can be acoustically distinguishable from those directed to subordinate individuals (e.g., Seyfarth and Cheney, 1984). Classification on basis of *Receiver-Rank* was orthogonal to sorting based on the *Infant* and *Move* variables, neither of which required that a given call be unambiguously associated with a particular receiver. Accordingly, the sample tested using the *Receiver-Rank* variable constituted a subset of the calls identified for analysis using the primary context variables.

### C. Apparatus

Calls were tape-recorded using Sennheiser directional microphones (ME80 heads with K3U power modules), Sony WC5M Professional Walkman cassette tape-recorders, and chromium-oxide metal tape cassettes. Tape-recorded grunts were subsequently digitized using the CSRE 4.0 software package (Avaaz Innovations, London, Ontario, Canada), implemented on a 486-type personal computer. Calls were sampled at 10.0 kHz using a DT2801A A/D board (Data Translation, Marborough, MA), a 16-channel audio interface, and 30-band, 1/4-octave, 20-kHz equalizer for low-pass filtering at approximately 4.8 kHz (preliminary analysis revealed only trace energy above 5.0 kHz). The bulk of acoustic analysis was conducted using the CSpeech version 4 software package (Paul Milenkovic, Dept. of Electrical and Computer Engineering, University of Wisconsin-Madison, 1415 Johnson Drive, Madison, WI 53706), also implemented on a 486-type personal computer. However, some measures were derived using the SIGNAL bioacoustics system (Engineering Design, Belmont, MA) on the same machine. The Number Cruncher Statistical System (NCSS; Jerry L. Hintze, 329 North 1000 East, Kaysville, UT 84037) was used for statistical analysis.

## D. Acoustic analysis

### 1. Call selection

Baboons typically produce grunts in readily identifiable bouts, consisting of a series of two or more calls given in rapid succession (although individual grunts are occasionally produced). Here, a bout was defined as a set of one or more grunts given by a single individual that was both preceded and followed by at least 10 s during which it did not grunt. The database consisted of a total of 1361 bouts of grunts recorded over the entire, yearlong observation period, all of which were digitized prior to further call selection. Nine-hundred and thirty-four of these bouts were produced by 16 adult females that could potentially serve as subjects in acoustic analyses. One grunt was randomly selected from each bout by a research assistant (if the bout consisted of only a single call, that call was selected). In each case, the assistant was aware of the age-sex class of the caller and its two-letter identity code, but was unaware of any other aspect of the circumstances of call production. Nine females, who accounted for a total of 464 of the recorded grunt bouts, were found to be represented by the largest number of calls produced in the greatest variety of social contexts, and were therefore designated as subjects.

Each call was then classified according to the context that had been noted for the corresponding bout during field observations, and the potential sample was then restricted to exemplars that were either Infant or Move related, as defined above. We included only calls that could be unambiguously assigned to either the Infant or Move contexts, for instance excluding vocalizations that were given to an infant while the group was simultaneously on the move. Following context-based classification, each of the nine female's calls were examined in the chronological order of their original occurrence to further screen out vocalizations showing recording-related acoustic artifacts. The calls eliminated were those whose onset and offset points were unmeasurably obscured by background noise and exemplars that included extraneous sounds. In some instances, digital high-pass filtering at 50 Hz was used to improve the quality of the calls. The signal-to-noise ratio (variable SNR) of each call was then measured by subtracting the average rms background noise level immediately preceding or following the vocalization from the average rms energy over the duration of the vocalization. The call-selection process continued until a final sample of 216 grunts had been identified. This set included 24 grunts recorded from each of the 9 subjects, with 12 calls from the Infant context and 12 calls from the Move context in each case.

### 2. $F_0$ -related measures

Three methods were used to extract the fundamental frequency of each call. First,  $F_0$  was determined from the call's time waveform using the CSpeech program (variable  $F_{0\_60\%}$ ). This algorithm is based on a least-mean-square, autocorrelation method described by Milenkovic (1987). It was applied over the middle 60% of each call, so as to maximize the probability of successful pitch extraction by excluding the lowest-amplitude cycles in each waveform. A repre-

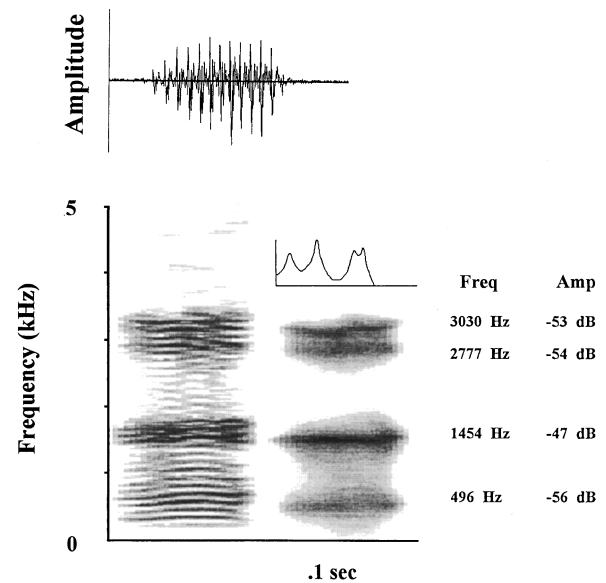


FIG. 2. A typical adult female chacma baboon grunt waveform (left) and corresponding digital spectrograms based on 256-point FFT (middle-left) and 14-coefficient, 256-point LPC computations (middle-right). A single LPC-based smoothed spectral envelope derived from the midpoint of the call is shown above the spectrograms; values listed on the right represent the frequency and amplitude values for each peak in this function. From Owren and Linker (1995), adapted with permission.

sentative baboon grunt waveform is shown in Fig. 2.  $F_0$  values derived using this approach were spot checked for each waveform through direct cursor measurement of one or more cycles.

The second approach, using the SIGNAL bioacoustics package, derived both the  $F_0$  and a number of other related measures from digital spectrograms. Each spectrogram was based on 2048-point FFTs computed using a Hanning analysis window, providing 9.8-Hz resolution. The cursor was used to measure frequency values of adjacent partials in the power spectrum, which was derived by centering the analysis window on the temporal midpoint of each call. Measurements were made of 3–5 partials in each call (as many as was practicable) in the portion of the frequency range showing greatest overall energy levels. These values were used to calculate a mean interval between the partials, which were hypothesized to be evenly spaced at multiples of the  $F_0$  (variable  $F_{0\_Harmonics}$ ).

The third approach also made use of the SIGNAL-derived digital spectrograms, but involved deriving a fundamental frequency contour directly from that representation. This approach (described in Owren and Casale, 1994) involved first drawing time and frequency boundaries around the visually identified fundamental frequency or a higher-amplitude partial using cursors or the mouse, and then automatically extracting the highest-amplitude frequency from within each column of the spectrogram's underlying matrix using an automated peak-detection function in SIGNAL. The mean overall  $F_0$  was then calculated from these contours (variable  $F_{0\_Contour}$ ).

### 3. Measurements from the time waveform

In addition to providing the basis for the  $F_{0\_60\%}$  variable described above, the time waveform was also used for measurement of overall call duration (variable  $Dur$ ) and SNR.

### 4. Spectral measurements

Frequency and amplitude measurements of four major spectral peaks were derived from spectral slices computed through covariance-based linear predictive coding [LPC; see Owren and Linker (1995) and Owren and Bernacki (in press) for reviews of LPC approaches and applications in bioacoustics]. For each call, a single slice was computed using 15 coefficients and a 512-point Hamming analysis window centered on the temporal midpoint of the sound. Each LPC spectrum was checked by superimposing it on a corresponding 512-point fast Fourier transform (FFT) and displaying these characterizations simultaneously with a spectrographic representation computed using a 300-Hz analysis filter bandwidth. In a small number of cases, the LPC function was recomputed with a slightly higher or lower coefficient number to provide a more accurate match to the FFT spectrum. Cursor-based measurements of center frequency (variables  $F1f-F4f$ ) and amplitude (variables  $F1a-F4a$ ) were made from the spectral peaks. In order to characterize frequency and amplitude relationships among the peaks, a series of relational measures were computed from these absolute values. Using values from all possible pairs of peaks, relationships were computed for both frequency (e.g., variable  $F2f-F1f$ ) and amplitude values (e.g., variable  $F1a-F2a$ ). Finally, a measure of the overall slope of the frequency spectrum (variable  $LPC\_Tilt$ ) was derived by exporting the spectral slice to a spreadsheet program and calculating its least-squares regression line.

## II. RESULTS

### A. General characteristics of the vocalization sample

The number of calls occurring in the 216 bouts from which the grunts were selected varied substantially ( $M = 4.38$  calls,  $s.d.=4.52$ , range 1–18). Using analysis of variance (ANOVA) with Context and Caller as factors, this variation was found to be statistically significant both between the Infant ( $M=641$ ,  $s.d.=5.31$ ) and Move contexts ( $M=2.34$ ,  $s.d.=2.54$ ),  $F(1, 198)=75.9$ ,  $p<0.001$ , and among the individual animals, whose means ranged from 2.33 to 7.75,  $F(8, 198)=7.0$ ,  $p<0.001$ . An interaction effect was also found,  $F(8, 198)=5.55$ ,  $p<0.001$ . Similar analyses of call durations revealed that the sounds were quite brief ( $M=109.4$  ms,  $s.d.=31.2$ ), showed no significant difference between the Infant ( $M=111.6$  ms,  $s.d.=33.2$ ) and Move ( $M=107.1$  ms,  $s.d.=29.0$ ) contexts,  $F(1, 198)=1.97$ ,  $p=0.16$ , but did show a main effect of Caller,  $F(8, 198)=19.4$ ,  $p<0.001$ ) and a significant interaction effect,  $F(8, 198)=2.69$ ,  $p<0.01$ .

Sixty-one calls were excluded during the preliminary call-selection process due to compromised recording quality, 28 in the Infant and 33 in the Move contexts, respectively.

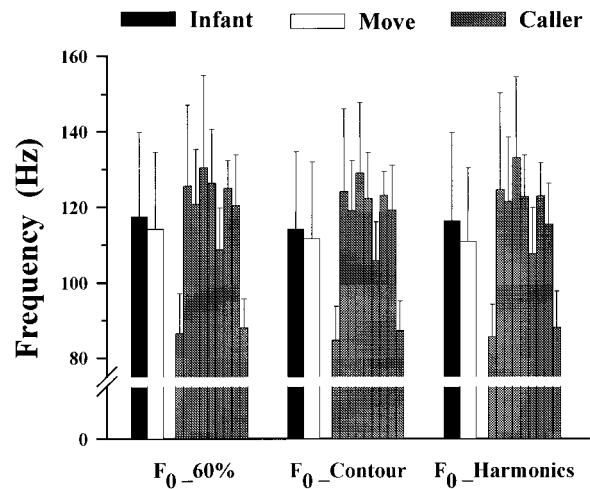


FIG. 3. Means and standard deviations of fundamental frequency values derived from the time waveform ( $F_{0\_60\%}$ ), first partial ( $F_{0\_Contour}$ ), and intervals between partials ( $F_{0\_Harmonics}$ ) are shown for the Infant and Move contexts (left) and for each of the nine adult female baboons (right).

The number of calls excluded did not differ by context,  $t(16)=0.19$ ,  $p=0.86$ , two-tailed, but did differ by individual caller ( $M=6.78$ ,  $s.d.=6.82$ ),  $\chi^2(8)=54.9$ ,  $p<0.001$ . The mean SNR value for calls included in the sample was 15.8 dBv (overall range 13.2–17.6 dBv). An ANOVA, using Context and Caller as factors, revealed no significant main effect of the former (Infant:  $M=15.2$ ,  $s.d.=6.26$ ; Move:  $M=16.3$ ,  $s.d.=5.94$ ),  $F(1, 197)=1.93$ ,  $p=0.17$ . However, both a main effect of Caller,  $F(8, 197)=2.05$ ,  $p=0.04$ , and an interaction effect,  $F(8, 197)=4.98$ ,  $p<0.001$ , were found. Follow-up, pairwise Scheffe comparisons of SNR values from individual animals revealed no significant differences. A separate test of mean SNR values from calls produced to subordinate ( $M=16.9$ ,  $s.d.=6.54$ ) versus dominant ( $M=14.8$ ,  $s.d.=5.74$ ) females was not statistically significant,  $t(109)=1.68$ ,  $p=0.10$ , two-tailed.

Based on the receiver identities noted in the field, the sample of 216 calls produced by the context- and acoustic-quality-related selection procedures was found to include 104 calls that had been given to either adult males, immature animals, or no apparent receiver, 72 calls given to subordinate females, and 40 calls given to dominant females. Analyses involving classification based on the Receiver-Rank variable described above thus included only the latter, 112 grunts.

### B. Fundamental frequency measures

Means and standard deviations for each of the three fundamental frequency measures are shown in Fig. 3, by Context and by Caller. A one-way ANOVA (that treated individual calls as subjects, thereby inflating its power) showed that these three sets of fundamental frequency measurements did not differ from one another,  $F(2, 639)=0.45$ ,  $p=0.64$ . Overall correlations among the three variables were then calculated, based on measurements from individual calls ( $n = 212$  calls, with 4 missing values). The resulting Spearman-

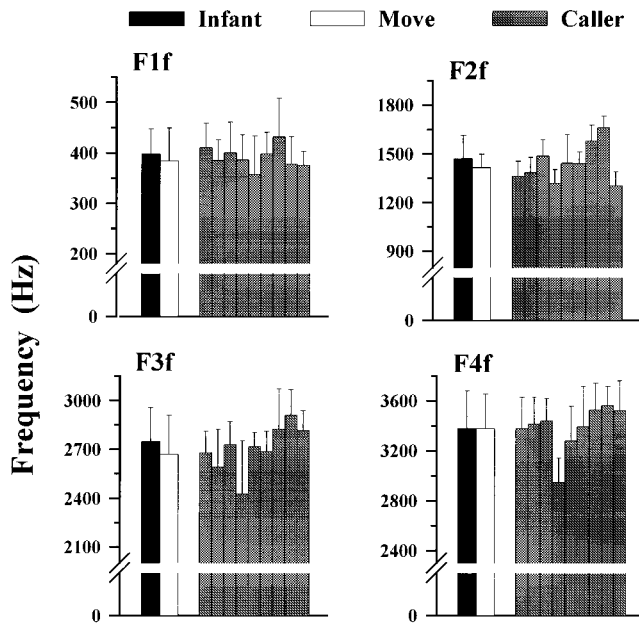


FIG. 4. Means and standard deviations of spectral peak frequencies. Overall values are shown for the Infant and Move contexts (left) and for each of the nine adult female baboons (right). The order of presentation for individual animals corresponds to that of Fig. 3 in each case.

rank correlation matrix revealed coefficients of 0.93, 0.95, and 0.96, comparing  $F_{0-60\%}$  with  $F_{0-Harmonics}$ ,  $F_{0-Harmonics}$  with  $F_{0-Contour}$ , and  $F_{0-60\%}$  with  $F_{0-Contour}$ , respectively. The Spearman-rank correlation test is more conservative than the corresponding, parametric Pearson  $r$  statistic (according to Hintze, 1995), and each of these correlations was statistically significant at a level of  $p < 0.001$ . To simplify further analyses,  $F_{0-60\%}$  alone was used as a representative measure of fundamental frequency, which will hereafter be referred to as  $F_0$ .

### C. Spectral peak frequencies and their relationships to fundamental frequency

Means and standard deviations for spectral peak frequencies are shown in Fig. 4, by Context and by Caller. Multiple regression analysis, in which  $F_0$  was used as a factor and  $F1f-F4f$  were entered as dependent variables, produced a statistically significant outcome,  $F(4, 179) = 5.50$ ,  $p < 0.001$ , which could be traced to individual relationships between  $F_0$  and  $F2f$ ,  $t = 3.54$ ,  $p < 0.001$ , and between  $F_0$  and  $F4f$ ,  $t = -3.08$ ,  $p < 0.01$ . However, the adjusted  $R$ -squared value of 0.09 observed for the overall relationship indicated that it was weak (this measure serves as a primary index of the equation's goodness-of-fit to the data—a value near 1.0 indicates a strong linear relationship, while a value near 0 indicates little or no association; see Hintze, 1995). Tested individually,  $F_0-F2f$  and  $F_0-F4f$   $R$ -squared values were 0.02 and 0.05, respectively. When each baboon was tested separately using this multiple regression approach, a statistically significant relationship was found between  $F_0$  and the four formant frequencies in only one case,  $F(4, 19) = 3.37$ ,  $p = 0.03$ . This outcome was associated with

TABLE I. Means and standard deviations of overall formant frequencies (Hz) with resulting estimates of vocal tract length (cm), based on idealized calculations for a straight, uniform tube closed at one end.

Formant	$M$	s.d.	Estimated vocal tract length (cm)
$F1$	437.7	21.4	19.7
$F2$	1442.1	121.3	18.0
$F3$	2706.3	147.5	15.9
$F4$	3374.6	201.0	17.9
Overall mean estimate			17.9

an adjusted  $R$ -squared value of 0.29. The overall mean adjusted  $R$ -squared value for the other eight individuals was 0.07.

### D. Vocal tract length estimates

Overall means for the first 4 formant frequencies (shown in Table I) were used to estimate the length of an adult female baboon's vocal tract. The estimates are based on modeling the vocal tract as a straight tube, closed at one end. The expected length ( $l$ ) in cm can be calculated as  $l = (2k + 1)c/4F$ , where  $F$  is the formant frequency,  $k$  is the formant number minus 1, and  $c$  is 34 400 cm/s, the speed of sound (see Lieberman and Blumstein, 1988). Individual estimates derived from each formant are also shown in the table, where the overall estimated length is 17.9 cm. An ANOVA was used to test the variation among the averaged length estimates derived from each animal's mean formant frequencies, and showed that significant variation did occur,  $F(3, 32) = 17.11$ ,  $p < 0.001$ . Scheffe tests indicated that this overall effect resulted from the estimates based on  $F1$  and  $F3$  frequencies. In these cases, the derived lengths were significantly longer and shorter, respectively, than expected (based on the mean overall estimate).

### E. Differences by context and caller

#### 1. Analysis of variance

Statistically significant differences were readily found on numerous dependent measures when Context and Caller were used as MANOVA factors, as shown in Table II. In these analyses, each variable was first subjected to  $z$ -score transformation to ensure normality, and three groupings of conceptually related acoustic variables were tested separately. As 28 of the calls did not exhibit an identifiable  $F4$ , all measures related to this formant were excluded from the analyses at this point. A skewing effect resulting from uneven distribution of the 28 missing cases across the Caller and Context variables was thereby avoided. The first grouping of the remaining variables (*Absolute-Spectral*) consisted of the absolute frequency ( $F1f-F3f$ ) and amplitude ( $F1a-F3a$ ) measures for each of the spectral peaks. The second grouping (*Relative-Spectral*) consisted of the relational measurements, including LPC\_Tilt values. The third grouping (*Waveform-Related*) consisted of measures related to the time waveform, namely SNR, Dur, and  $F_0$ .

In each case in which a statistically significant context-related difference occurred, the direction of this difference

TABLE II. Overall MANOVA results using Context and Caller as factors in conjunction with three variable groupings—Absolute-Spectral (AS), Relative-Spectral (RS), and Waveform-Related (Wv). Single and double asterisks denote statistical significance at the 0.05 and 0.01 levels, respectively.

Grouping	Factor	df	F value	Statistically significant variables
AS	Context	6,182	7.73**	$F1f^*, F2f^{**}, F2a^{**}, F3f^{**}, F3a^{**}$
	Caller	48,900	8.38**	$F1f^{**}, F2f^{**}, F2a^{**}, F3f^{**}, F3a^{**}$
	Interaction	48,900	1.89*	$F1a^{**}, F2f^{**}, F3f^*$
RS	Context	7,181	7.59**	$F2f-F1f^{**}, F3f-F1f^{**}, F3f-F2f^{**}, F2a-F3a^{**}, LPC\_Tilt^{**}$
	Caller	56,980	6.04**	$F2f-F1f^{**}, F3f-F1f^{**}, F3f-F2f^{**}, F2a-F3a^{**}, LPC\_Tilt^{**}$
	Interaction	56,980	1.41*	$F2f-F1f^{**}, F3f-F1f^*, F3f-F2f^*, F1a-F2a^{**}, F1a-F3a^{**}$
Wv	Context	1,3	4.84**	$F_{0\_60\%}^{**}$
	Caller	8,24	13.01**	$Dur^{**}, F_{0\_60\%}^{**}$
	Interaction	8,24	2.71**	$Dur^*, F_{0\_60\%}^{**}$

was tallied by subject. In most cases, statistical significance for a given variable did not imply a consistent pattern of underlying mean differences for individual animals. For  $F2f$ , however, a nonparametric one-proportion test did reveal statistically significant consistency in direction of difference (i.e., eight of nine cases showed a higher mean value in the Infant context  $M = 1470.2$ ,  $s.d. = 148.5$ , than in the Move context,  $M = 1415.3$ ,  $s.d. = 146.1$ ,  $z = 2.33$ ,  $p = 0.02$ ), while  $LPC\_Tilt$  closely approached statistically significant consistency (seven of nine cases showed a higher mean value in the Infant context,  $M = -6.17$ ,  $s.d. = 131$ , than in the Move context,  $M = -5.25$ ,  $s.d. = 1.93$ ,  $z = 1.67$ ,  $p = 0.10$ ).

Receiver-Rank was tested with ANOVA using the same three variable groupings. A significant overall outcome that emerged for Absolute-Spectral measures,  $F(6, 99) = 2.81$ ,  $p = 0.01$ , was found to be entirely due to the value of  $F2f$ . Overall, this frequency was significantly lower when the call was produced to a subordinate animal ( $M = 1383$ ,  $s.d. = 111.7$ ) than to a dominant individual ( $M = 1468$ ,  $s.d. = 111.3$ ),  $F(1, 104) = 14.2$ ,  $p < 0.001$ . However, this outcome did not occur consistently among the various subjects. Relative-Spectral measures also revealed a significant difference by Receiver-Rank,  $F(7, 98) = 2.16$ ,  $p = 0.045$ , which was due to the variable  $F2f-F1f$ ,  $F(1, 104) = 14.2$ ,  $p < 0.001$ . As no differences were found in  $F1f$  for either grouping,  $F2f$  can be presumed to underly the significant effect in  $F2f-F1f$ . Again, no evidence of consistency in the direction of difference emerged among the various animals. No significant differences were found for Waveform-Related measures,  $F(3, 107) = 0.99$ ,  $p = 0.40$ .

## 2. Discrimination function analyses

The overall lack of consistency among individuals in the direction of differences when producing calls in the three social contexts was borne out by discriminant function analyses conducted separately for Context, Caller, and Receiver-Rank. In these analyses (again excluding  $F4$ -related measures), classification was based on manual stepwise entry of variables that continued until we had entered all variables whose associated  $F$ -value produced a  $p$ -value of less than 0.10. Results are illustrated in Table III, including the measures entered in each case. The 8 of 15 variables entered for Caller classification was a significantly higher proportion than the 2 of 15 variables (i.e.,  $F2f$  and  $LPC\_Tilt$ ) entered for Context classification ( $z = -2.32$ ,

$p = 0.02$ , nonparametric two-proportions test), or the 1 of 15 variables (i.e.,  $F2f$ ) entered for Receiver-Rank classification ( $z = -2.78$ ,  $p < 0.01$ ). Similarly, error reduction when calls were sorted by Caller was significantly greater than when sorted by Context ( $z = 7.23$ ,  $p < 0.001$ ) or by Receiver-Rank ( $z = 4.61$ ,  $p < 0.001$ ).

## 3. Relationships between $F2f$ and spectral tilt measures

$F2f$  and  $LPC\_Tilt$  were found to be significantly correlated,  $t = 3.17$ ,  $p < 0.01$ , but with an adjusted  $R$ -square value of only 0.045. Analogously, testing the relationship of  $F2f$  to the relational amplitude variables in multiple regression, an overall significant effect,  $F(2, 202) = 5.10$ ,  $p = 0.01$ , was associated with an adjusted  $R$ -square value of only 0.039.  $LPC\_Tilt$  was, however, found to be significantly correlated with the various relational amplitude measures when tested with them individually (the latter were themselves significantly intercorrelated). As a result of the strong overall relationship found,  $F(2, 202) = 67.5$ ,  $p$

TABLE III. Variables entered in discriminant function analysis for classification of calls by Context, Caller, and Receiver-Rank, with resulting classification accuracy. Table entries show the order in which each variable was entered and its associated  $F$ -to-enter value.

Variable	Context	Caller	Receiver-Rank
Dur		(3) 16.8	
$F_{0\_60\%}$		(2) 28.1	
$F1f$		(8) 2.7	
$F1a$		(7) 4.6	
$F2f$	(2) 13.6	(1) 30.2	(1) 14.3
$F2a$			
$F3f$			
$F3a$		(5) 3.1	
$F2f-F1f$			
$F3f-F1f$		(4) 8.1	
$F3f-F2f$			
$F1a-F2a$			
$F1a-F3a$			
$F2a-F3a$			
$LPC\_Tilt$	(1) 16.5	(6) 3.3	
Classification accuracy	Context	Caller	Receiver-Rank
By chance	50.0%	11.1%	50.0%
Statistically	62.8%	65.7%	66.7%
Error reduction	25.6%	58.6%	33.3%

< 0.001, with an adjusted *R*-square value of 0.395, LPC\_Tilt and the relational amplitude variables were considered to be approximately interchangeable measures of overall spectral tilt.

### III. DISCUSSION

#### A. Baboon vocal production mechanisms

##### 1. Source and filter

As the baboon sound production system anatomically consists of vocal folds and a supralaryngeal vocal tract, the apparent role of vocal-fold vibration was examined by testing the relationship between the evident fundamental frequency of this vibration and energy found elsewhere in the sound spectrum. As primate calls can pose difficulties in automated vocal pitch extraction, for instance due to mismatches between algorithm design and the acoustic properties of particular calls (see Owren and Linker, 1995), three conceptually distinct measures were used for this aspect of the analysis.  $F_0$ \_60% values were based on a time-domain method whose combination of short-term autocorrelation and center clipping (i.e., setting low-amplitude samples to zero) is designed to attenuate all but the longest waveform periodicities (see Hess, 1983; O'Shaughnessy, 1987). This approach produced results that were indistinguishable from measurements of the lowest-frequency partial in the frequency spectrum,  $F_0$ \_Contour, which in turn were a virtual match to values based on the frequency intervals observed between adjacent partials at higher frequencies,  $F_0$ \_Harmonics. This outcome indicates that the primary source energy of grunt calls is periodic vibration, which can be presumed to occur in the vocal folds.

Direct tests of the independence of the source energy and subsequent vocal tract filter were not possible with these data, so an indirect approach of examining correlations between  $F_0$  values and spectral frequency peak values was used instead. No relationships were observed between  $F_0$  and formant peaks  $F1$  and  $F3$ . For  $F2$  and  $F4$ , statistically significant correlations were found, but the degree of relationship was very small in each case. A significant relationship emerged for only one of the nine animals when each was tested individually. While the degree of relationship was more convincing in that one case, the overall dataset demonstrates independence between  $F_0$  values and spectral peak frequencies, leading us to conclude that these measures reflect separable components of the sound-production system.

These data do not, of course, provide evidence concerning the spectrum of the source energy function itself. Such information would be necessary to gauge the relative contributions of frequency-dependent amplitude values and vocal tract resonance effects in shaping the overall amplitude of a given harmonic. In addition, baboons exhibit supra-laryngeal air sacs (see Schön Ybarra, 1995), whose possible role in grunt production cannot be evaluated at this point. Nonetheless, the occurrence of humanlike, formant-related spectral effects is strongly indicated.

These results are consistent with findings of both Zhinkin (1963) and Andrew (1976), who examined spectral

energy peaks in grunts recorded from *hamadryas*, and *hamadryas* and *anubis* baboons, respectively. Testing the peaks using a 32-band spectrometer, Zhinkin found them to generally match the formant frequencies of his own vocal imitations of the sounds. Andrew (1976), who examined the calls of one juvenile male and one infant using a sound spectrograph, also found evidence of a formant-based vocal tract transfer function imposed on a glottally related harmonic structure.

##### 2. Comparisons to human speech production

The mean  $F_0$  of 114.5 Hz (99.9–128.2 Hz interquartile range, based on the  $F_0$ \_60% measure) associated with these adult female baboon grunt calls is very similar to the typical spontaneous speaking fundamental frequency of 123 Hz reported for young human adult males (see Baken, 1987). Baboon grunts are often produced in the absence of obvious arousal and are also lower pitched than are calls like *screams* or *fear barks* that are typically associated with such arousal (Seyfarth and Cheney, 1997). This apparent match between baboons and humans in the fundamental frequency of normative vocalizations indicates between-species similarity in the length of the vibrating component of the vocal folds [see Titze (1994) for a review of vocal-fold mechanics]. Adult male baboons typically produce grunts with even lower fundamental frequencies (Seyfarth and Cheney, 1997) and are therefore likely to have longer vocal folds than do either adult female baboons or humans.

Mean frequency values for the first three formants, which were 438, 1442, and 2706 Hz, respectively, can be closely matched to typical formant frequencies reported for human adult males by Peterson and Barney (1952; only three formants are typically reported for vowels). The estimated vocal tract length of 17.9 cm derived from the calls is consequently similar to the frequently cited characterization of the adult human male vocal tract as being 17.5 cm long. Species in the *Papio* genus all exhibit elongated muzzles (e.g., Hill, 1972; Napier and Napier, 1985) with requisitely long mandibles, oral, and nasal cavities (as illustrated in Fig. 1). A chacma baboon skull measured by Elliot (1913) was reported to have a palatal length of 19.0 cm and mandible length of 15.8 cm. Given that the larynx is set just behind and below the oral cavity (e.g., Negus, 1949), the estimated vocal tract lengths derived here are consistent with these measurements.

Figure 5 shows the mean  $F1$  and  $F2$  values measured in the baboons superimposed on Peterson and Barney's well-known mapping of the American-English vowel space, based on 76 men, women, and children. As is apparent both from the means and from the larger shaded area representing all the data points, the variation in these formant values covers a significant, but centrally located proportion of this space. Overall, this outcome is consistent with Lieberman and colleagues' hypothesis that the typical "vowel space" of primates exhibits a restricted range of formant variation that directly reflects the relatively inflexible nature of their "single-tube" vocal tracts. However, the formant values we measured indicate that the adult female baboon vocal tract is

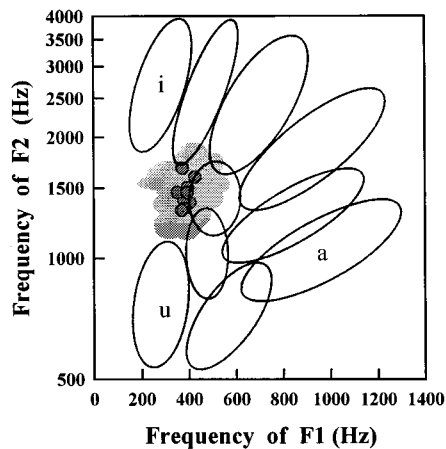


FIG. 5. Mean frequency measurements for  $F_1$  and  $F_2$  from the nine adult female baboons (shaded circles) are shown in the American-English  $F_2/F_1$  vowel-space mapping of Peterson and Barney (1952). The larger shaded area represents the range of variation found among all 216 calls examined.

not entirely uniform over its length and can therefore not be exactly matched by an idealized straight-tube resonator. In addition, further investigation of baboon calls may reveal instances of greater formant variability than was found here, although due to their low pitch and dense harmonic structure, grunts do appear to be the most likely vocalization in the baboon repertoire to exhibit clear, and functionally important formant patterns.

As discussed in greater detail below, the data provided only modest indications of changes in formant-related spectral peak patterns by social context. This outcome is somewhat surprising, in that primates appear capable of a number of articulatory movements that modify the vocal tract transfer function during the production process (see Lieberman, 1975). In fact, the work of Hauser and colleagues (Hauser, 1992; Hauser *et al.*, 1993; Hauser and Schön Ybarra, 1994), indicates that rhesus monkeys routinely do alter the shapes and lengths of their vocal tracts through various articulatory maneuvers, thereby influencing resonant-frequency outcomes. A similar outcome was indirectly indicated by laryngographic measurements involving two Sykes's monkey subjects tested by Brown and Cannito (1995).

While some data specifically relevant to baboons are available from previous work, they shed little additional light on this question of functional modification of vocal tract positioning. Zhinkin (1963), for instance, reported successful x-ray photography of one subject during vocalization and found no evidence of significant tongue movement (specifically, "pharyngeal narrowing"). However, the sound in question was a high-pitched squeal, a call type that is unlikely to be usefully modified by flexible articulator positioning due to having a high fundamental. In contrast, Andrew (1976) specifically reported that one of his young baboon subjects routinely produced changes in formant position through mandible and tongue movements. Unfortunately, the maneuvers (lip-smacking and tongue protrusion) were ones that are not typical of normative call production in adult baboons (Seyfarth and Cheney, 1997).

In sum, then, the data from both the current study and

past work lend little support to the notion that changes in formant position are a critical aspect of baboon vocal communication. However, a firm conclusion must await more comprehensive analyses of various calls in this species, specifically including work with captive animals. The nature of both the source energy and the supralaryngeal vocal tract resonances involved in baboon sound production would be amenable to separate examination using any of a number of the approaches developed in the context of articulatory models of human speech (reviewed by Rubin and Vatikiotis-Bateson, in press). Techniques like laryngography (Brown and Cannito, 1995) and blocking of articulatory control through anesthetic injection (Hauser and Schön Ybarra, 1994) have already been successfully applied to other monkey species.

## B. Acoustic variation related to individual identity

### 1. Functional significance of variation and the importance of formants

Discrimination of kin from nonkin, as well as individuals within both of these categories, has been proposed to be of overriding significance both to the evolution of social behavior in general (Hamilton, 1964; Trivers, 1971) and in the kin-biased behaviors of primates specifically (e.g., Cheney *et al.*, 1986; Smuts *et al.*, 1987). A number of studies of monkeys and apes have shown either individual variability in the acoustic structure of vocalizations, discrimination among individual conspecific callers, or both [reviewed in Snowdon (1986); see also Gouzoules *et al.*, 1986; Gouzoules and Gouzoules, 1990; Owren *et al.*, 1992; Rendall *et al.*, 1996]. While no systematic, vocalization-based evidence is available concerning discrimination or recognition of individuals or kin in baboons, the requisite acoustic cues to support these expected functions appear to be present in grunt calls [see also Cheney *et al.* (1995a, b) for indirect evidence of recognition of individually distinctive grunts, screams, and fear barks in these animals].

Acoustic variation among individuals was found for a variety of the measures used here, of which ten variables were found to be useful predictors of caller identity in discriminant function analysis. While most of the acoustic measures were related to some aspect of the spectral characteristics of these calls, variables  $Dur$  and  $F_0$  were also found to be important in sorting calls, at least by individual. However, it should be noted that any given animal is also likely to be able to facultatively alter both of these attributes when grunting, as evidenced by the ability of these baboons to produce other, longer or much higher-pitched, tonal call types (Seyfarth and Cheney, 1997). It is therefore requisitely unlikely that either call duration, vocal pitch, or any combination of the two can serve as the primary cues to individual identity in this vocal repertoire.

In contrast, frequency and amplitude characteristics of the spectral peaks found in the grunts emerged as a predominant source of identity-based classificatory power that can also be expected to be stable both within and among callers. Based on the source-filter principles of vocal production discussed in the Introduction, we hypothesize that these spectral characteristics act as markers of small-scale variation in for-

mant frequencies and amplitudes that in turn result from minor anatomical differences in the structures and cavities of the vocal tracts of individual baboons. Very few spectral peak characteristics appeared to vary among the social contexts examined using our sample. In conjunction with the brief nature of these grunts and an evident lack of within-call formant changes noted during the acoustic analysis process, this evidence suggests to us that vocal tract resonance patterns such as those imprinted on these harmonically rich grunts may provide a particularly stable and dependable source of information marking the identity of the calling animal across a variety of vocalization contexts [discussed by Rendall *et al.* (in review) based on rhesus monkey vocalizations].

## 2. Perceptual salience of the observed variation in acoustic features

From a perceptual point of view, the acoustic attributes indexed by the measures found to be of greatest interest in statistical analysis also can in general be expected to be salient to listening baboons, although some of the data supporting this view are derived from other primate species. As a group, Old World monkeys appear to be generally similar in overall sensitivity to pure-tone stimuli across the audible frequency range [e.g., Stebbins (1973), but also see Brown and Waser (1984); Owren *et al.* (1988)], including chacma baboons tested by Hienz *et al.* (1982). In comparison to humans, these animals are somewhat less sensitive to frequencies below 500 Hz, significantly more sensitive to frequencies above 8000 Hz, and comparable (albeit slightly less sensitive) in the intermediate range. Thus, the acoustic energy characterized here should be readily perceptible to our subjects.

While macaques and some other species have been found to exhibit six to ten times higher frequency-difference thresholds than humans when both are tested with intermediate-frequency, pure-tone stimuli (see, for instance, Sinnott and Brown, 1993; Sinnott *et al.*, 1985; Sinnott *et al.*, 1987a), the degree of variation in  $F_0$  values between individual animals found here is greater than expected difference threshold values would be. Monkeys have also been found to be quite sensitive to variation in duration, at least with respect to brief sounds (Sinnott *et al.*, 1987b). Finally, based on work by Sommers *et al.* (1992), one can expect that frequency variation within and between the formants demonstrated in the grunts is perceptually salient to the baboons (see Table I). In this study, Japanese macaques tested with single- and multi-formant tone complexes showed formant-frequency discrimination abilities at 500 Hz (e.g., 12.5 Hz) and 1.4 kHz (e.g., 22.8 Hz) that were comparable to those of humans. This finding is consistent with the demonstrations, noted earlier, that both chacma baboons (Hienz and Brady, 1988) and Japanese macaques (Sinnott, 1989) readily discriminate among typical American-English vowel sounds. Both species perform least well when presented with pairs of sounds that human listeners are most likely to confuse, while doing best when the vowel pairs are the ones most readily distinguished by humans. Although direct evidence is not available, one can expect as an associated capability that

baboons are also sensitive to variation in formant amplitude and general characteristics of spectral shape, like tilt.

## C. Acoustic variation related to social context

When grunting to an infant or to a mother with an infant, all adult females produced longer call bouts (approximately six grunts per bout) than when grunting while on a move (approximately two grunts per bout). This result suggests that, from the baboons' perspective, these two social contexts were markedly different. However, when we tested for acoustic differences in grunts produced in the two circumstances, there was little evidence of systematic acoustic variation. While ANOVA-based tests revealed statistically significant differences between Infant and Move contexts for a number of variables, concomitant, nonsystematic variation was also likely to be found among individual callers.

In only two cases, namely variables  $F2f$  and LPC\_Tilt, did the animals in the sample show possibly persuasive consistency in the direction of statistically significant context-related mean differences. These two measures were also the only variables found to be useful in discriminant-function-based call classification. Proportionately, classification accuracy was significantly lower by context and relative dominance rank than by individual caller. Interpreting both LPC\_Tilt and the relative amplitude variables broadly, as measures of overall spectral shape, the results indicated a more steeply tilted spectrum when calls were produced in the Infant context and a shallower spectrum when vocalizations occurred in the Move context. Testing Receiver-Rank,  $F2f$  was found to be the only important predictor variable in discriminant-function analysis. Due to an overall lack of correlation between tilt-related and  $F2f$  values, the two variables may represent functionally independent aspects of social communication in these calls. Overall, the former appears most likely to be related to behavioral context regardless of caller identity or relative rank of the interactants, while the latter is more clearly associated with the identity-related attributes of sender, receiver, or both.

## D. Recording-related considerations

As the recordings were made in the field from free-ranging subjects, it is appropriate to consider whether inadvertent but inevitable variation in recording conditions may have affected the results for these often quiet, usually low-arousal calls. On the one hand, as the study animals were part of a baboon group that has been under continuous observation for many years, calls were routinely recorded from close quarters and were therefore of very high quality by the standards of primate fieldwork (cf. Owren and Bernacki, 1988). On the other hand, at least two sources of recording-related variation may have differentially affected the analysis outcomes for various measures. These possibly confounding factors were the distance of the observer from the vocalizing animal and variation in the angle of the microphone relative to the direction of call emission. The former is important due to the greater relative attenuation of high-frequency energy over distance, while the latter is related to the relative directionality of high-frequency energy when broadcast from a

vocal-tract opening [e.g., Brown (1989); see Gerhardt (in press) for a general review of these issues in field recording]. In both cases, then, any circumstantial differences in distance or microphone angle could be expected to create corresponding variability in the relative energy of higher frequencies.

However, with the exception of overall tilt of the spectrum, the variables that were of greatest interest in the analyses (i.e., those related to absolute formant frequency, fundamental frequency, and call duration) are likely to have been quite robust to incidental variation in recording conditions. For instance, as transmission of formant-frequency information represented in a tonal complex depends only on preservation of the relative amplitudes of adjacent harmonics, this aspect of the calls is unlikely to have been affected by global, frequency-dependent effects. Likewise, both fundamental frequency and overall call duration characteristics are arguably unaffected by selective attenuation of high-frequency energy. Spectral tilt, in contrast, directly reflects the relative amplitude of high- and low-frequency energy and may therefore have been more vulnerable to recording-related variation.

#### IV. CONCLUSIONS

Overall, the data considered here demonstrate a significant acoustic resemblance between grunts produced by chacma baboons and prototypical, centrally located vowel sounds produced by humans. This resemblance is likely to be rooted in a general cross-species similarity in vocal folds and supralaryngeal vocal tract, whose actions are most similar for humans and baboons when the former produce sounds using relaxed, neutral articulator positioning. However, the formant-frequency characteristics of the grunts we measured also indicate that a uniform, straight tube does not provide an entirely adequate model for understanding the evident resonance effects associated with the vocal tracts of adult female baboons.

Differences in grunt acoustics were clearly related to the identity of particular calling animals. Based on statistical analysis of call features, probable asymmetries in the degree of volitional control over various aspects of vocal production, and the demonstrated perceptual salience of spectral sound features (in baboons and Japanese macaques), variability in formant-related attributes is likely to be a predominant source of biologically significant, acoustically based differentiation among grunts. Variation in spectral peaks, shown to be independent of variation in  $F_0$ , was associated both with individual caller and the social context of the vocalization event. In the former case, individual variation in details of vocal tract anatomy appeared to be reflected in a variety of acoustic measures. In the latter case, frequency of the second formant and overall spectral tilt appear to be differentially associated with the Infant and Move contexts, and with the relative rank of the animal to whom the call is directed, respectively.

However, acoustic differences across social contexts were by no means as clear as were acoustic differences among individuals, and the context-based differences were not equally strong for all animals tested. In conjunction with the observation that spectral-tilt-related measures were likely

to be the most susceptible to incidental recording effects, the results indicate that while adult female baboons may have shown differences in the acoustic features of their grunts when they were vocalizing to individuals that were more dominant or more subordinate, they are most likely to have used acoustically distinct grunt types in the Infant and Move contexts. Research is currently underway to test this hypothesis through playback experiments conducted with the study group and to examine how the distance and angle of recording affect the acoustic features of tape-recorded calls.

#### ACKNOWLEDGMENTS

We thank the Office of the President and the Department of Wildlife and National Parks of the Republic of Botswana for permission to conduct research in the Moremi Game Reserve. We are grateful to W. Hamilton, R. Boyd, Mongabe Kgosiekae, and Mokopi Mokopi for logistical support, to Eric Bokenkamp for selecting, editing, and analyzing grunts, and to Bridget Sly for further assistance in acoustic analysis. Drawings shown in Fig. 1 were generously created and donated by Michael Graham. Interpretation of the results benefited from valuable discussions with Drew Rendall and JoAnne Bachorowski. Two anonymous reviewers and Tecumseh Fitch provided helpful comments on a previous version of this work. Portions of the current results were presented at the November 1995 meeting of the Acoustical Society of America. Recording of vocalizations was supported by grants from the National Geographic Society, the Research Foundation of the University of Pennsylvania, the Institute for Cognitive Science at the University of Pennsylvania, and National Institutes of Health Grant No. HD-29433. Equipment and funding for assistance in vocalization analysis was provided by the College of Liberal Arts and Sciences at the University of Colorado at Denver, by a Division of Sciences Grant at the University of Otago, and the Dean's Development Fund at Reed College. Correspondence concerning this article should be addressed to Michael J. Owren, Department of Psychology, Reed College, 3203 SE Woodstock Blvd., Portland, OR 97202. Electronic mail may be sent via the internet to michael.owren@reed.edu

Andrew, R. J. (1976). "Use of formants in the grunts of baboons and other nonhuman primates," in *Origins and evolution of language and speech* (Special Issue), edited by S. R. Harnad, H. D. Steklis, and J. Lancaster, Ann. (N.Y.) Acad. Sci. **280**, 673–693.

Baken, R. J. (1987). *Clinical Measurement of Speech and Voice* (College-Hill, Boston).

Bolwig, N. (1959). "A study of the behaviour of the chacma baboon, *Papio ursinus*," *Behaviour* **14**, 136–163.

Brown, C. H. (1989). "The measurement of vocal amplitude and vocal radiation pattern in blue monkeys and grey-cheeked mangabeys," *Bioacoustics* **1**, 253–271.

Brown, C. H., and Cannito, M. P. (1995). "Modes of vocal variation in Sykes's monkey (*Cercopithecus albogularis*) squeals," *J. Comp. Psych.* **109**, 398–415.

Brown, C. H., Gomez, R., and Waser, P. M. (1995). "Old World monkey vocalizations: Adaptation to the local habitat?" *Animal Beh.* **50**, 945–961.

Brown, C. H., and Waser, P. M. (1984). "Hearing and communication in blue monkeys (*Cercopithecus mitis*)," *Animal Beh.* **32**, 66–75.

Bulger, J., and Hamilton, W. J. (1988). "Inbreeding and reproductive success in a natural chacma baboon, *Papio cynocephalus ursinus*, population," *Animal Beh.* **36**, 574–578.

- Byrne, R. W. (1981). "Distance vocalizations of Guinea baboons (*Papio papio*) in Senegal: An analysis of function," *Behaviour* **78**, 283–313.
- Cheney, D. L., and Seyfarth, R. M. (1982). "How vervets perceive their grunts: Field playback experiments," *Animal Beh.* **30**, 739–751.
- Cheney, D. L., and Seyfarth, R. M. (in press). "Reconciliatory grunts by dominant female baboons influence victims behaviour," *Animal Beh.*
- Cheney, D. L., Seyfarth, R. M., and Palombit, R. M. (1996). "The function and mechanisms underlying baboon 'contact' barks," *Animal Beh.* **52**, 507–518.
- Cheney, D. L., Seyfarth, R. M., and Silk, J. B. (1995a). "The role of grunts in reconciling opponents and facilitating interactions among adult female baboons," *Animal Beh.* **50**, 249–257.
- Cheney, D. L., Seyfarth, R. M., and Silk, J. B. (1995b). "The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: Evidence for causal reasoning?" *J. Comp. Psych.* **109**, 134–141.
- Cheney, D. L., Seyfarth, R. M., and Smuts, B. (1986). "Social relationships and social cognition in nonhuman primates," *Science* **234**, 1361–1366.
- Chiba, T., and Kajiyama, J. (1941). *The vowel: Its Nature and Structure* (Tokyo-Kaiseikan, Tokyo).
- Clark, A., and Wrangham, R. (1993). "Acoustic analysis of wild chimpanzee pant hoots: Do Kibale forest chimpanzees have an acoustically distinct food arrival pant hoot?" *Am. J. Prim.* **31**, 99–109.
- Elliot, D. G. (1913). *A Review of the Primates, Vol. 2* (American Museum of Natural History, New York).
- Fant, G. (1960). *The Acoustic Theory of Speech Production* (Mouton, The Hague).
- Fitch, W. T., and Hauser, M. D. (1995). "Vocal production in nonhuman primates: Acoustics, physiology, and functional constraints on 'honest' advertisement," *Am. J. Prim.* **37**, 191–219.
- Gautier, J.-P., and Gautier, A. (1977). "Communication in Old World monkeys," in *How Animals Communicate*, edited by T. A. Sebeok (Indiana U.P., Bloomington, IN), pp. 890–964.
- Geist, F. D. (1933). "Nasal cavity, larynx, mouth and pharynx," in *The Anatomy of the Rhesus Monkey*, edited by C. G. Hartman and W. L. Straus, Jr. (Hafner, New York), pp. 189–209.
- Gerhardt, C. (in press). "Acoustic signals of animals: Recordings, field measurements, analysis, and description," in *Animal Acoustic Communication: Sound Analysis and Research, Methods*, edited by S. L. Hopp, M. J. Owren, and C. S. Evans (Springer-Verlag, Heidelberg).
- Gouzoules, H., and Gouzoules, S. (1990). "Matrilineal signatures in the recruitment screams of pigtail macaques, *Macaca nemestrina*," *Behaviour* **115**, 327–347.
- Gouzoules, H., Gouzoules, S., and Marler, P. (1986). "Vocal communication: A vehicle for studying social relationships," in *The Cayo Santiago Macaques: History, Behavior, and Biology*, edited by R. G. Rawlins and M. J. Kessler (SUNY, Albany, New York), pp. 111–129.
- Green, S. (1975). "Variation of vocal pattern with social situation in the Japanese monkey (*Macaca fuscata*): A field study," in *Primate Behavior, Vol. 4*, edited by L. A. Rosenblum (Academic, New York), pp. 1–102.
- Grimm, R. J. (1967). "Catalogue of sounds of the pigtailed macaque (*Macaca nemestrina*)," *J. Zool. Soc. London* **152**, 361–373.
- Hamilton, W. D. (1964). "The genetical evolution of social behavior. I, II," *J. Theor. Biol.* **7**, 1–52.
- Hauser, M. D. (1992). "Articulatory and social factors influence the acoustic structure of rhesus monkey vocalizations: A learned mode of production?" *J. Acoust. Soc. Am.* **91**, 2175–2179.
- Hamilton III, W. J., and Arrowood, P. C. (1978). "Copulatory vocalizations of chacma baboons (*Papio ursinus*), gibbons (*Hylobates hoolock*), and humans," *Science* **200**, 1405–1409.
- Hamilton, W. J., and Bulger, J. (1992). "Facultative expression of behavioral differences between one-male and multimale savanna baboon groups," *Am. J. Prim.* **28**, 61–71.
- Hauser, M. D., and Fowler, C. A. (1991). "Declination in fundamental frequency is not unique to human speech: Evidence from nonhuman primates," *J. Acoust. Soc. Am.* **91**, 363–369.
- Hauser, M. D., and Schön Ybarra, M. (1994). "The role of lip configuration in monkey vocalizations: Experiments using xylocaine as a nerve block," *Brain Lang.* **46**, 232–244.
- Hauser, M. D., Evans, C. S., and Marler, P. (1993). "The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalizations," *Animal Beh.* **45**, 423–433.
- Hess, W. (1983). *Pitch Determination of Speech Signals* (Springer-Verlag, Heidelberg).
- Hienz, R. D., and Brady, J. V. (1988). "The acquisition of vowel discriminations by nonhuman primates," *J. Acoust. Soc. Am.* **84**, 186–194.
- Hienz, R. D., Turkkan, J. S., and Harris, A. H. (1982). "Pure tone thresholds in the yellow baboon (*Papio cynocephalus*)," *Hearing Res.* **8**, 71–75.
- Hill, W. C. O. (1972). *Evolutionary Biology of the Primates* (Academic, New York).
- Hill, W. C. O., and Booth, A. H. (1957). "Voice and larynx in African and Asiatic *Colobidae*," *Bombay Nat. Hist. Soc.* **54**, 309–321.
- Hilloowala, R. A. (1975). "Comparative anatomical study of the primate hyoid apparatus," *Am. J. Anat.* **142**, 367–385.
- Hintze, J. (1995). *NCSS 6.0 User's Guide* (Number Cruncher Statistical Systems, Kaysville, UT).
- Hohmann, G. M. (1989). "Vocal communication of wild bonnet macaques (*Macaca radiata*)," *Primates* **30**, 325–345.
- Hohmann, G. M., and Herzog, M. O. (1985). "Vocal communication in lion-tailed macaques (*Macaca silenus*)," *Folia Primatol.* **45**, 148–178.
- Hopp, S. L., Owren, M. J., and Evans, C. S., Eds. (in press). *Animal Acoustic Communication: Sound Analysis and Research Methods* (Springer-Verlag, Heidelberg).
- Kelemen, G. (1938). "Comparative anatomical studies on the junction of larynx and resonant tubes," *Acta Oto-Laryngol.* **26**, 276–283.
- Kelemen, G. (1949). "The anatomical basis of phonation in the chimpanzee," *J. Morphol.* **82**, 229–256.
- Kelemen, G. (1969). "Anatomy of the larynx and the anatomical basis of vocal performance," *Chimpanzee* **1**, 165–186.
- Lieberman, P. (1968). "Primate vocalizations and human linguistic ability," *J. Acoust. Soc. Am.* **44**, 1574–1584.
- Lieberman, P. (1969). "On the acoustic analysis of primate vocalizations," *Behav. Res. Methods Instrum. Comput.* **5**, 169–174.
- Lieberman, P. (1975). *On the Origins of Language: An Introduction to the Evolution of Human Speech* (Macmillan, New York).
- Lieberman, P., and Blumstein, S. E. (1988). *Speech Physiology, Speech Production, and Acoustic Phonetics* (Cambridge U. P., Cambridge, MA).
- Lieberman, P., Crelin, E. S., and Klatt, D. H. (1972). "Phonetic ability and related anatomy of the newborn, adult human, Neanderthal man, and the chimpanzee," *Am. Anthropol.* **74**, 287–307.
- Lieberman, P., Klatt, D. H., and Wilson, W. A. (1969). "Vocal tract limitation on the vowel repertoires of rhesus monkey and other nonhuman primates," *Science* **164**, 1185–1187.
- Masataka, N., and Thierry, B. (1993). "Vocal communication of Tonkean macaques in confined environments," *Primates* **34**, 169–180.
- Milenkovic, P. (1987). "Least squares measures of voice perturbation," *J. Speech Hear. Res.* **30**, 529–538.
- Napier, J. R., and Napier, P. H. (1985). *The Natural History of the Primates* (MIT, Cambridge, MA).
- Negus V. E. (1949). *The Comparative Anatomy and Physiology of the Larynx* (Hafner, New York).
- O'Connell, S. M., and Cowlshaw, G. (1994). "Infanticide avoidance, sperm competition, and mate choice: The function of copulation calls in female baboons," *Animal Beh.* **48**, 687–694.
- O'Shaughnessy, D. (1987). *Speech Communication* (Addison-Wesley, Reading, MA).
- Owren, M. J. (1990). "Acoustic classification of alarm calls by vervets (*Cercopithecus aethiops*) and humans: II. Synthetic calls," *J. Comp. Psych.* **104**, 29–40.
- Owren, M. J., and Bernacki, R. H. (1988). "The acoustic features of alarm calls in vervet monkeys (*Cercopithecus aethiops*)," *J. Acoust. Soc. Am.* **5**, 1927–1935.
- Owren, M. J., and Bernacki, R. H. (in press). "Applying linear predictive coding (LPC) to frequency spectrum analysis of animal acoustic signals," in *Animal Acoustic Communication: Sound Analysis and Research Methods*, edited by S. L. Hopp, M. J. Owren, and C. S. Evans (Springer-Verlag, Heidelberg).
- Owren, M. J., and Casale, T. M. (1994). "Variations in fundamental frequency peak position in Japanese macaque (*Macaca fuscata*) 'coo' calls," *J. Comp. Psych.* **108**, 291–297.
- Owren, M. J., and Linker, C. D. (1995). "Some analysis methods that may be useful to acoustic primatologists," in *Current Topics in Primate Vocal Communication*, edited by E. Zimmermann, J. D. Newman, and U. Jürgen (Plenum, New York), pp. 1–27.
- Owren, M. J., Linker, C. D., and Rowe, M. P. (1993). "Acoustic features of tonal 'grunt' calls in baboons," *Acoustical Society of America biannual meeting*, Denver, CO.
- Owren, M. J., Seyfarth, R. M., and Cheney, D. L. (1995). "Acoustic indices of production mechanisms underlying tonal 'grunt' calls in baboons,"

- Acoustical Society of America biannual meeting, St. Louis, MO.
- Owren, M. J., Dieter, J. A., Seyfarth, R. M., and Cheney, D. L. (1992). "Food calls produced by adult female rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) macaques, their normally-raised offspring, and offspring cross-fostered between species," *Behaviour* **120**, 218–231.
- Owren, M. J., Hopp, S. L., Sinnott, J. M., and Petersen, M. R. (1988). "Absolute auditory thresholds in three Old World monkey species (*Cercopithecus aethiops*, *C. neglectus*, *Macaca fuscata*) and humans," *J. Comp. Psych.* **102**, 99–107.
- Peters, E. H. (1983). "Vocal communication in an introduced colony of feral rhesus monkey (*Macaca mulatta*)," Doctoral Dissertation, University of Florida, Gainesville.
- Peterson, G. E., and Barney, H. L. (1952). "Control methods used in a study of the identification of vowels," *J. Acoust. Soc. Am.* **24**, 175–184.
- Rendall, D., Owren, M. J., and Rodman, P. S., "The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations," *J. Acoust. Soc. Am.* (in review).
- Rendall, D., Rodman, P. S., and Emond, R. E. (1996). "Vocal recognition of individuals and kin in free-ranging rhesus monkeys," *Animal Beh.* **51**, 1007–1015.
- Richman, B. (1976). "Some vocal distinctive features used by gelada monkeys," *J. Acoust. Soc. Am.* **60**, 718–724.
- Rubin, P., and Vatikiotis-Bateson, E. (in press). "Measuring and modeling speech production," in *Animal Acoustic Communication: Sound Analysis and Research Methods*, edited by S. L. Hopp, M. J. Owren, and C. S. Evans (Springer-Verlag, Heidelberg).
- Schön Ybarra, M. (1995). "A comparative approach to the nonhuman primate vocal tract: Implications for sound production," in *Current Topics in Primate Vocal Communication*, edited by E. Zimmermann, J. D. Newman, and U. Jürgens (Plenum, New York), pp. 185–198.
- Seyfarth, R. M., and Cheney, D. L. (1997). "Vocal behavior in chacma baboons," (unpublished data).
- Seyfarth, R. M., and Cheney, D. L. (1984). "The acoustic features of vervet monkey grunts," *J. Acoust. Soc. Am.* **75**, 1623–1628.
- Seyfarth, R. M., Cheney, D. L., Harcourt, A. H., and Stewart, K. (1994). "The acoustic features of gorilla double-grunts and their relation to behavior," *Am. J. Prim.* **33**, 31–50.
- Sinnott, J. M. (1989). "Detection and discrimination of synthetic English vowels by Old World monkeys (*Cercopithecus*, *Macaca*) and humans," *J. Acoust. Soc. Am.* **86**, 557–565.
- Sinnott, J. M., and Brown, C. H. (1993). "Effects of varying signal and noise levels on pure-tone frequency discrimination in humans and monkeys," *J. Acoust. Soc. Am.* **93**, 1535–1540.
- Sinnott, J. M., Owren, M. J., and Petersen, M. R. (1987a). "Auditory frequency discrimination in primates: Species differences (*Cercopithecus*, *Macaca*, *Homo*) and humans," *J. Comp. Psych.* **101**, 126–131.
- Sinnott, J. M., Owren, M. J., and Petersen, M. R. (1987b). "Auditory duration discrimination in Old World monkeys (*Macaca*, *Cercopithecus*) and humans," *J. Acoust. Soc. Am.* **82**, 465–470.
- Sinnott, J. M., Petersen, M. R., and Hopp, S. L. (1985). "Frequency and intensity discrimination in humans and monkeys," *J. Acoust. Soc. Am.* **78**, 1977–1985.
- Smuts, B. B., Cheney, D. L., Seyfarth, R. M., Wrangham, R. W., and Struhsaker, T. T. (1987). *Primate Societies* (University of Chicago, Chicago).
- Snowdon, C. T. (1986). "Vocal communication," in *Comparative Primate Biology, Volume 2A: Behavior, Conservation, and Ecology*, edited by G. Mitchell and J. Erwin (Liss, New York), pp. 495–530.
- Sommers, M. S., Moody, D. B., and Prosen, C. A. (1992). "Formant frequency discrimination by Japanese macaques (*Macaca fuscata*)," *J. Acoust. Soc. Am.* **91**, 3499–3510.
- Stebbins, W. C. (1973). "Hearing of Old World monkeys (*Cercopithecinae*)," *Am. J. Phys. Anthropol.* **38**, 357–364.
- Stevens, K. N., and House, A. S. (1955). "Development of a quantitative description of vowel articulation," *J. Acoust. Soc. Am.* **27**, 484–493.
- Titze, I. R. (1994). *Principles of Voice Production* (Prentice-Hall, Englewood Cliffs, NJ).
- Trivers, R. (1971). "The evolution of reciprocal altruism," *Q. Rev. Biol.* **46**, 35–57.
- Waser, P. M. (1982). "The evolution of male loud calls in mangabeys and baboons," in *Primate Communication*, edited by C. T. Snowdon, C. H. Brown, and M. R. Petersen (Cambridge U. P., New York), pp. 117–143.
- Zhinkin, N. I. (1963). "An application of the theory of algorithms to the study of animal speech," in *Acoustic Behaviour of Animals*, edited by R. G. Busnel (Elsevier, Amsterdam), pp. 132–180.
- Zimmermann, E. (1981). "First record of ultrasound in two prosimian species," *Naturwissenschaften* **68**, 531.